

Space for WHAM: a multi-region, multi-stock generalization of the  
Woods Hole Assessment Model with an application to black sea  
bass:  
Supplementary Materials

Timothy J. Miller<sup>1,2</sup>

Kiersten L. Curti<sup>1,3</sup>

Alexander C. Hansell<sup>1,4</sup>

1 Northeast Fisheries Science Center, National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543, USA

2 orcid: 0000-0003-1411-1206, [timothy.j.miller@noaa.gov](mailto:timothy.j.miller@noaa.gov)

3 orcid: 0009-0009-1676-7269

4 orcid: 0000-0001-7827-1749

## Deriving the prior distribution for movement parameters

The Working Group fit a Stock Synthesis model (Methot and Wetzel 2013) that included tagging data with 2 seasons (6 months each) and 2 regions where a proportion  $\mu_1^*$  of the northern component moves to the south in one season and some proportion  $\mu_{2 \rightarrow 1}^*$  move back to the south in the second season (NEFSC 2023). The seasonal movement matrices for each season are

$$\boldsymbol{\mu}_1^* = \begin{bmatrix} 1 - \mu_{1 \rightarrow 2}^* & \mu_{1 \rightarrow 2}^* \\ 0 & 1 \end{bmatrix}$$

and

$$\boldsymbol{\mu}_2^* = \begin{bmatrix} 1 & 0 \\ \mu_{2 \rightarrow 1}^* & 1 - \mu_{2 \rightarrow 1}^* \end{bmatrix}.$$

To obtain estimates of movement proportions for the monthly intervals in the WHAM model, the half-year movement matrices were converted to monthly movement matrices by taking the root  $z_k$  of  $\boldsymbol{\mu}_k^*$  which are defined by the number of months of movement for each season (5 and 4, respectively). The roots of the matrices are calculated using an eigen decomposition of the matrices

$$\boldsymbol{\mu}_k = (\boldsymbol{\mu}_k^*)^{z_k} = \mathbf{V}_k \mathbf{D}_k^{z_k} \mathbf{V}_k^{-1},$$

where  $z_1 = 1/5$  for and  $z_2 = 1/4$ , and  $\mathbf{V}_k$  and  $\mathbf{D}_k$  are the matrix of eigenvectors (columnwise) and the diagonal matrix of corresponding eigenvalues of  $\boldsymbol{\mu}_k^*$ . The Working Group used a parametric bootstrap approach to determine an appropriate standard deviation for the prior distribution for the movement parameters. Stock Synthesis also estimates parameters on a transformed scale, but different from WHAM:

$$\mu_{r \rightarrow r'}^* = \frac{1}{1 + 2e^{-x_{r \rightarrow r'}}}$$

The estimated parameters and standard errors from the Stock Synthesis model were  $x_{1 \rightarrow 2} = -1.44$  and  $x_{2 \rightarrow 1} = 1.94$  and  $SE(x_{1 \rightarrow 2}) = 0.21$  and  $SE(x_{2 \rightarrow 1}) = 0.37$ . The resulting in the estimated proportions were  $\mu_{1 \rightarrow 2}^* = 0.11$  and  $\mu_{2 \rightarrow 1}^* = 0.78$ .

In WHAM, an additive logit transformation is used which is simply a logit transformation when there are only two regions:

$$\mu_{r \rightarrow r'} = \frac{1}{1 + e^{-y_{r \rightarrow r'}}}.$$

We simulated 1000 values from a normal distribution with mean and standard deviation defined by the parameter estimate and standard error  $\tilde{x}_{r \rightarrow r', b} \sim N(x_{r \rightarrow r'}, SE(x_{r \rightarrow r'}))$  from the Stock Synthesis model. For each simulated value we constructed  $\tilde{\boldsymbol{\mu}}_{r \rightarrow r', b}^*$ , took the appropriate root and calculated inverse logit for  $\tilde{y}_{r \rightarrow r', b}$ . We calculated the mean and standard deviation of the values  $y_{i, b}$ . The mean values did not differ meaningfully from the transformation of the original estimates ( $y_{1 \rightarrow 2} = -3.79$  and  $y_{2 \rightarrow 1} = -0.79$ ) and the standard deviation was approximately 0.2 for both parameters.

## Bottom temperature anomalies

The Working Group created bottom temperature observations from a high resolution ocean bottom temperature product by du Pontavice et al. (2023). The annual observations for each region are defined by the average over all spatial bottom temperature values for February and March by region and year. Similarly, the Working Group calculated standard errors from the standard deviation of all values in the region and the total number of values for a given year. We created regional bottom temperature anomalies by subtracting means for each region across all years.

## Diagnostics

### Jitter fits for model $M_0$

WHAM by default completes three newton steps after the `stats::nlminb` minimization function completes to reduce the gradient at the minimized NLL. However, this generally has negligible effects on model estimates and the NLL. To reduce computation time, we did not complete these newton steps when performing jitter fits of the model. Without the Newton steps, the maximum (absolute) gradient sizes are generally less than 0.01 for models that converge satisfactorily.

The 50 jitter fits demonstrated that a local minimum was obtained for the original fit of model  $M_0$  (Figure S12). One lower NLL was obtained with unacceptable gradients (No. 25), but a slightly lower NLL was found with a satisfactory gradient for 3 of the jitters (Nos. 9, 13, 29). However, one of the jitter fits (No. 9) did not provide a non-zero estimate of the variance parameter for one of the indices and the other two provided identical results and we refit model  $M_0$  and all remaining models using the better parameter estimates as initial values.

### Jitter fits for model $M_1$

The 50 jitter fits gave no evidence of a better minimization of the NLL. Three lower NLLs were obtained, but with unacceptably large gradients (Figure S13). The largest differences in parameter estimates for these three jitters were for numbers at age and selectivity random effects variance and correlation parameters.

### Self test for model $M_1$

Initial fits to simulated data from model  $M_1$  showed estimation of the observation error standard deviation multiplier for the recreational catch-per angler indices in the north and south regions was unstable. Many of the fits to the simulated data produced implausible estimates at the 0 boundary for these parameters (very negative values on log-scale). However, across all fits including those with poor convergence, estimation of SSB and fishing mortality was reliable (Figure S1). We also fit analogous models with the multiplier parameters fixed at the true values, which did improve convergence, but larger bias was estimated for fishing mortality and SSB for the northern component.

## References

- du Pontavice, H., Chen, Z., and Saba, V.S. 2023. A high-resolution ocean bottom temperature product for the northeast U.S. Continental shelf marine ecosystem. *Progress in Oceanography* **210**: 102948. doi:10.1016/j.pocean.2022.102948.
- Methot, R.D., and Wetzel, C.R. 2013. Stock Synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research* **142**(1): 86–99. doi:10.1016/j.fishres.2012.10.012.
- NEFSC. 2023. Report of the black sea bass (*Centropristis striata*) research track stock assessment working group. Available at [https://www.mafmc.org/s/a\\_2023\\_BSB\\_UNIT\\_RTWG\\_Report\\_V2\\_12\\_2\\_2023-1.pdf](https://www.mafmc.org/s/a_2023_BSB_UNIT_RTWG_Report_V2_12_2_2023-1.pdf).

Table S1: Configuration of age composition likelihoods, mean selectivity models, and selectivity random effects models for each age composition data component. For all logistic-normal likelihoods, any ages observed as zeros are treated as missing.

Data component	Age Composition Likelihood	Mean Selectivity model	Random effects Model
North commercial fleet	Dirichlet-Multinomial	age-specific (ages > 3 fully selected)	AR1 correlation by age and year
North recreational fleet	Logistic-normal (Independent)	age-specific (ages > 6 fully selected)	AR1 correlation by age and year
South commercial fleet	Logistic-normal (AR1 correlation)	logistic	None
South recreational fleet	Logistic-normal (AR1 correlation)	logistic	None
North recreational CPA index	Logistic-normal (Independent)	age-specific (ages > 1 fully selected)	AR1 correlation by year
North VAST index	Dirichlet-Multinomial	age-specific (ages > 4 fully selected)	AR1 correlation by age and year
South recreational CPA index	Logistic-normal (AR1 correlation)	age-specific (ages > 2 fully selected)	None
South VAST index	Logistic-normal (AR1 correlation)	age-specific (ages > 1 fully selected)	None

Table S2: Model AIC weights for each retrospective peel.

Model	Peel							
	0	1	2	3	4	5	6	7
$M_0$	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01
$M_1$	0.45	0.43	0.42	0.38	0.41	0.37	0.36	0.38
$M_2$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$M_3$	0.30	0.31	0.25	0.23	0.24	0.23	0.25	0.23
$M_4$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$M_5$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$M_6$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$M_7$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$M_8$	0.15	0.15	0.20	0.24	0.21	0.24	0.22	0.23
$M_9$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$M_{10}$	0.10	0.11	0.12	0.14	0.13	0.15	0.15	0.14
$M_{11}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$M_{12}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$M_{13}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

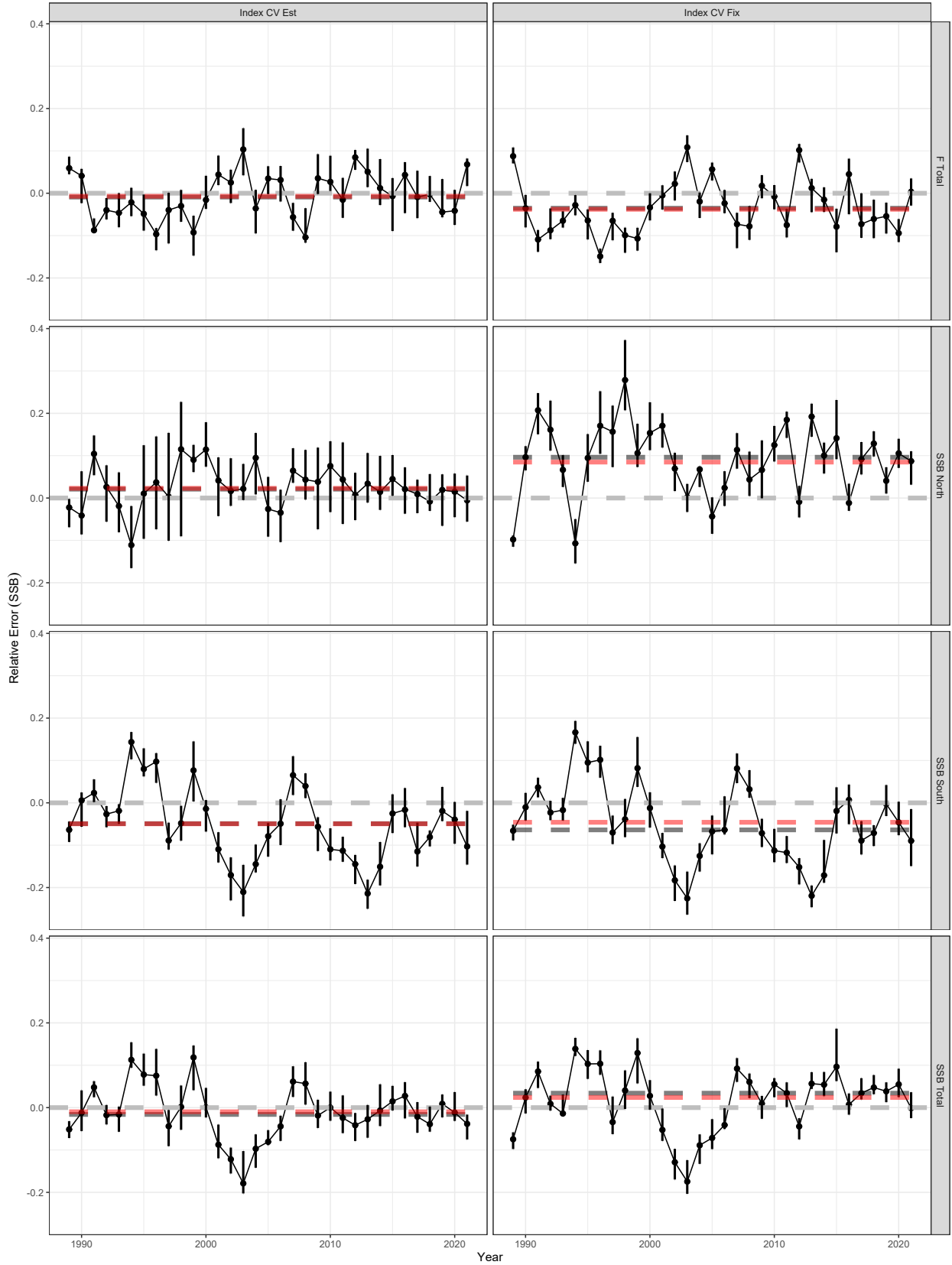


Figure S1: Median relative error of SSB (Total and by stock component) and total fully-selected fishing mortality for estimation models fitted to simulated data from model  $M_1$  where the observation variance of log-indices are fixed and estimated. Black and Red dashed lines represent the median of the annual medians and the median across all annual relative errors, respectively. Vertical lines represent 95% confidence intervals.

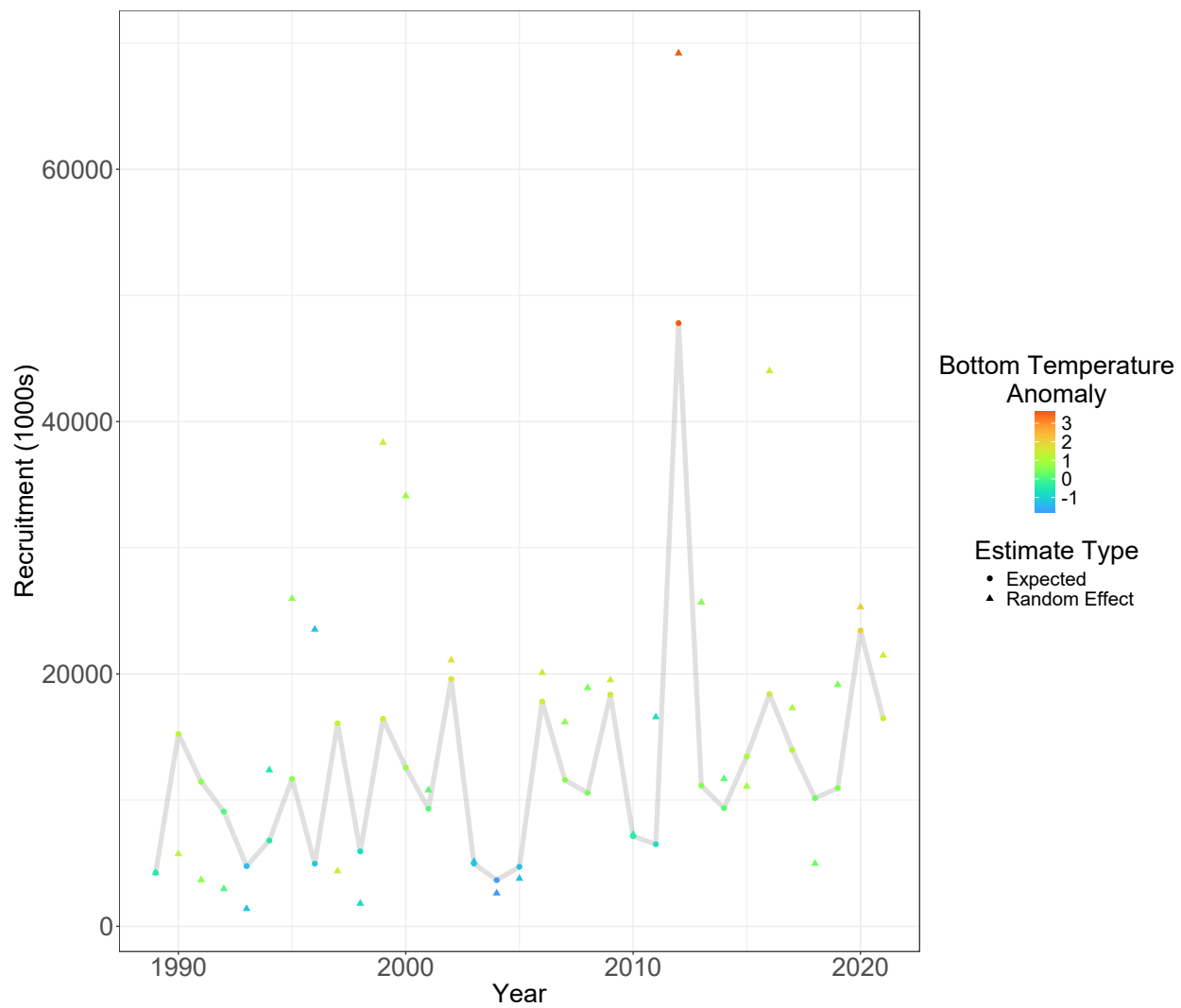


Figure S2: Expected and random effect recruitment estimates for the northern stock component. Color of points defined by the corresponding annual bottom temperature anomaly.

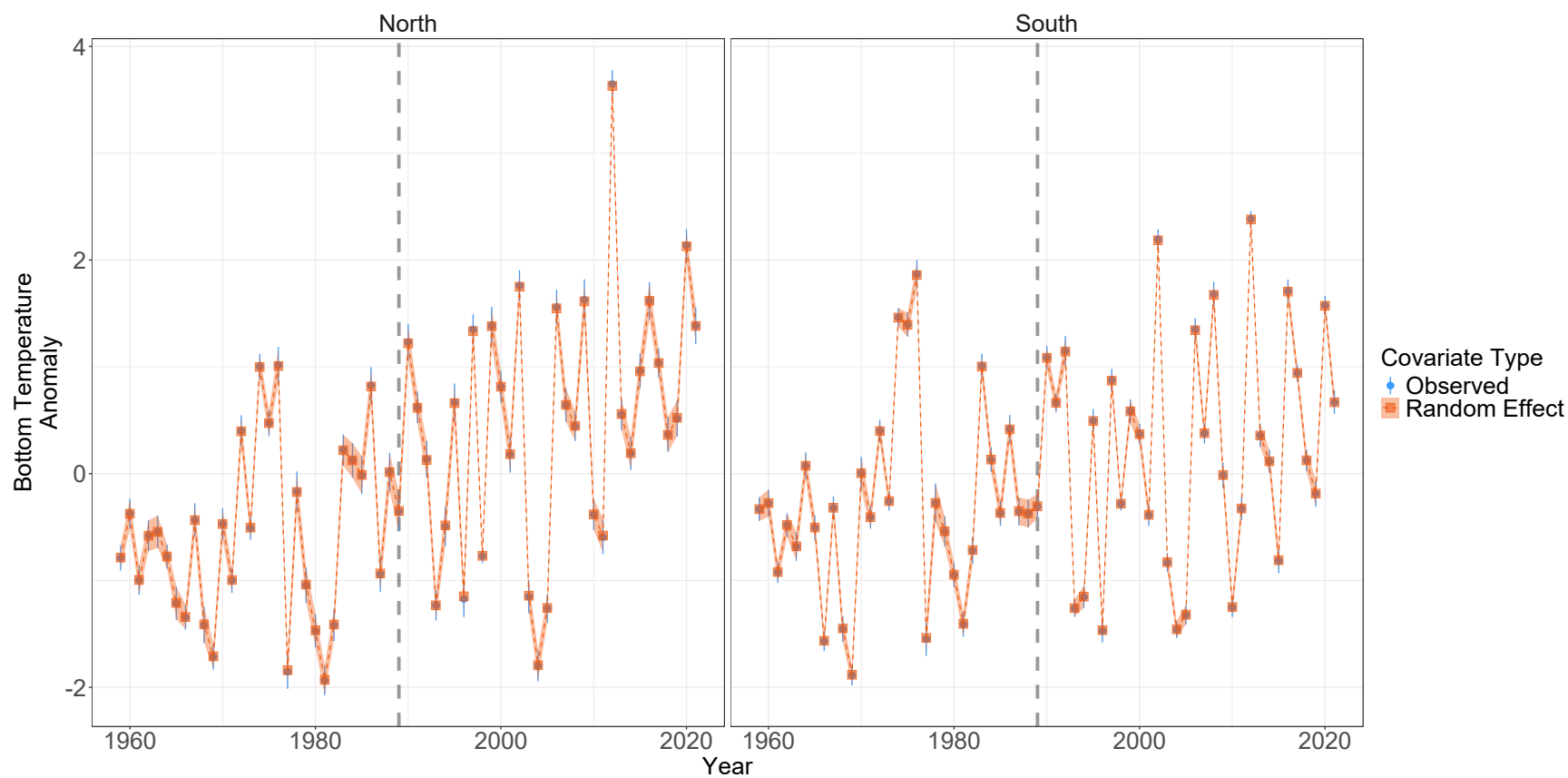


Figure S3: Observations with 95% confidence intervals (points with vertical lines) and posterior estimates with 95% confidence intervals (lines with polygons) of bottom temperature anomalies in the north and south regions from model  $M_1$ . Gray vertical line defines the first year that the black sea bass stock is modeled.



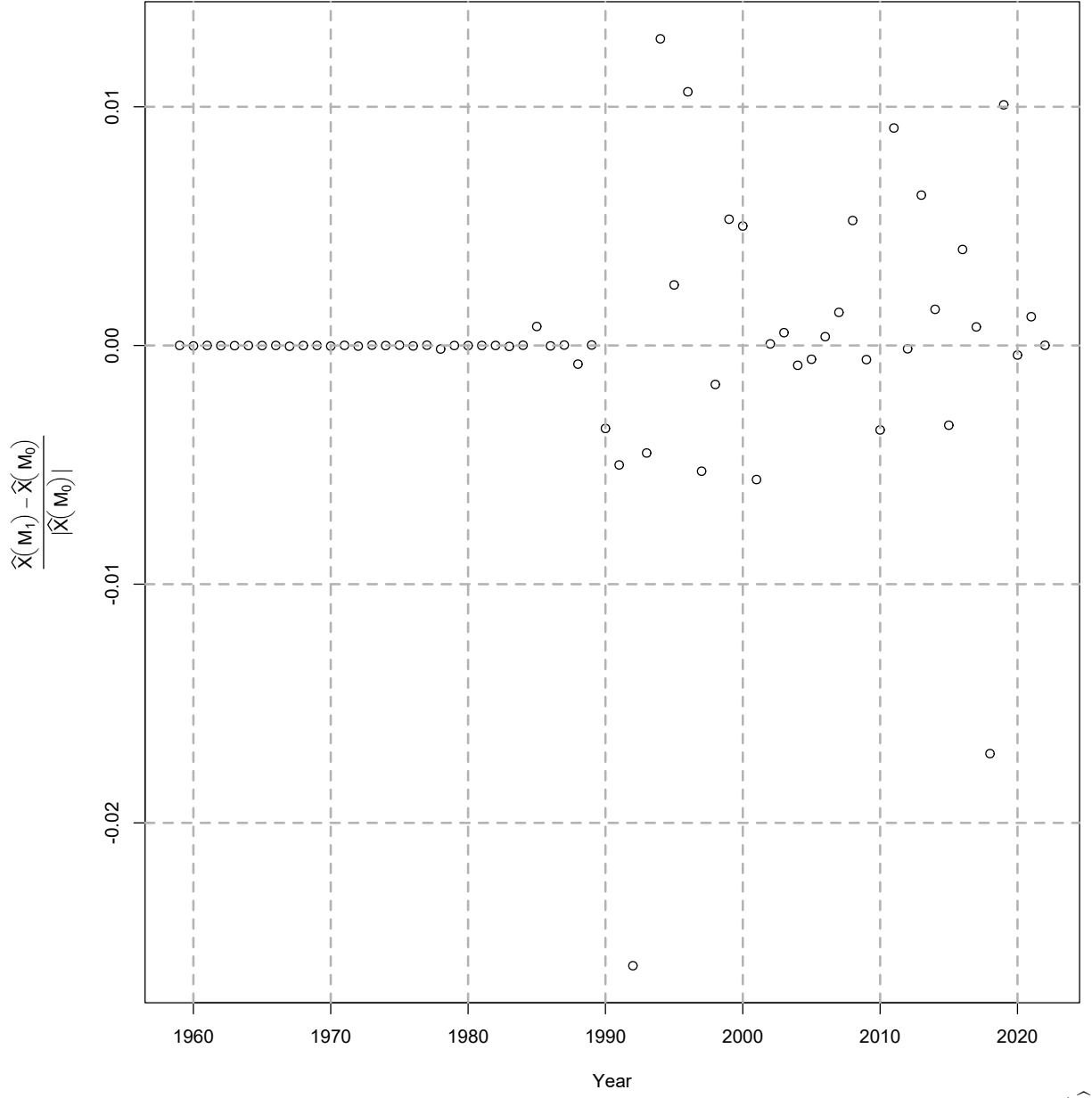


Figure S4: Relative differences in posterior estimates of northern region bottom temperature anomalies ( $\hat{X}$ ) from the null model without effects on recruitment ( $M_0$ ) and with effects on the northern stock component ( $M_1$ ).

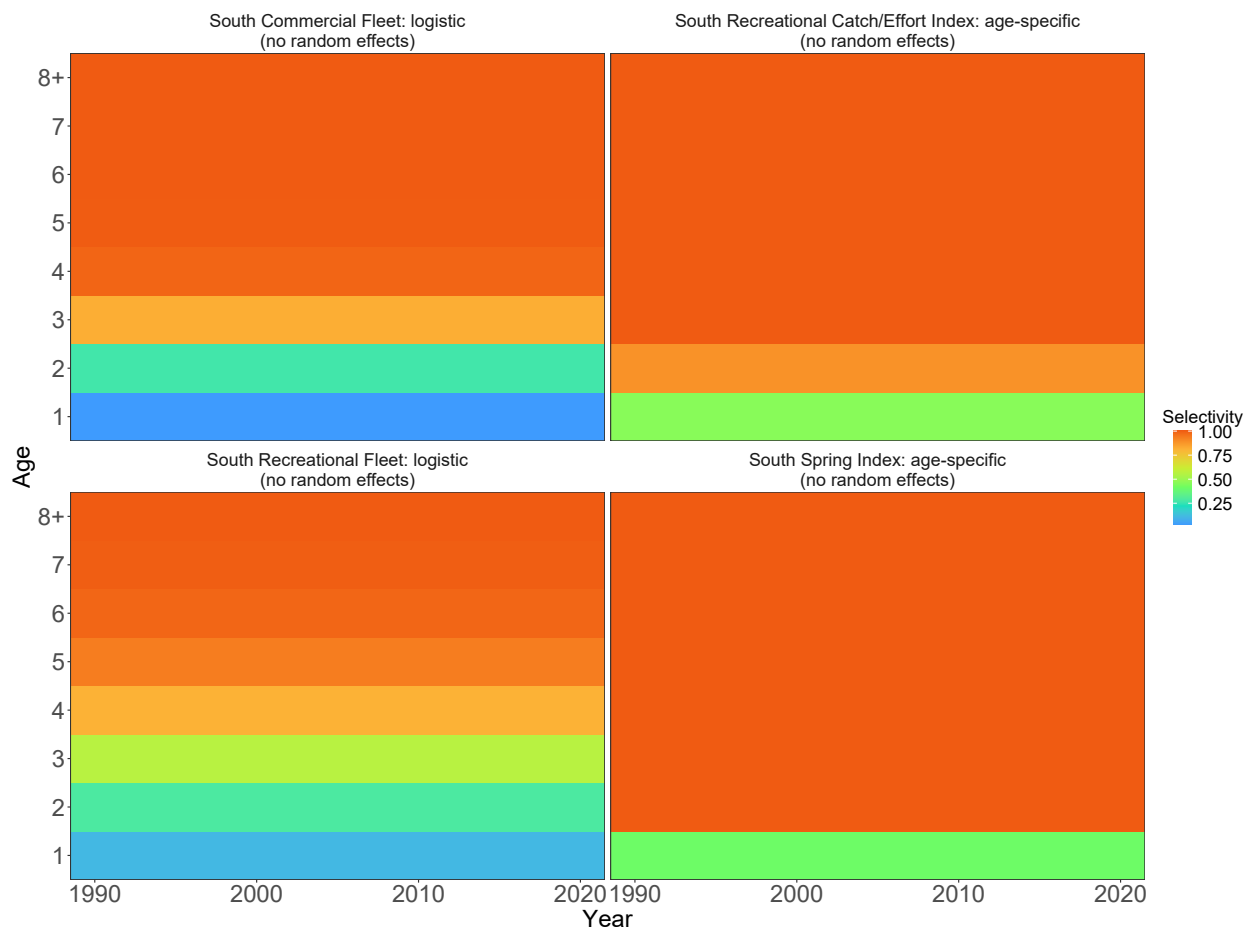


Figure S5: Selectivity for fleets and indices in the southern region.



Figure S6: Estimates of annual SSB and fishing mortality rates for the best performing model  $M_1$  and models that are otherwise the same except where 1) movement rates are fixed at the means for the prior distribution, 2) a constant natural mortality rate is estimated, or 3) there is no movement for either stock component.

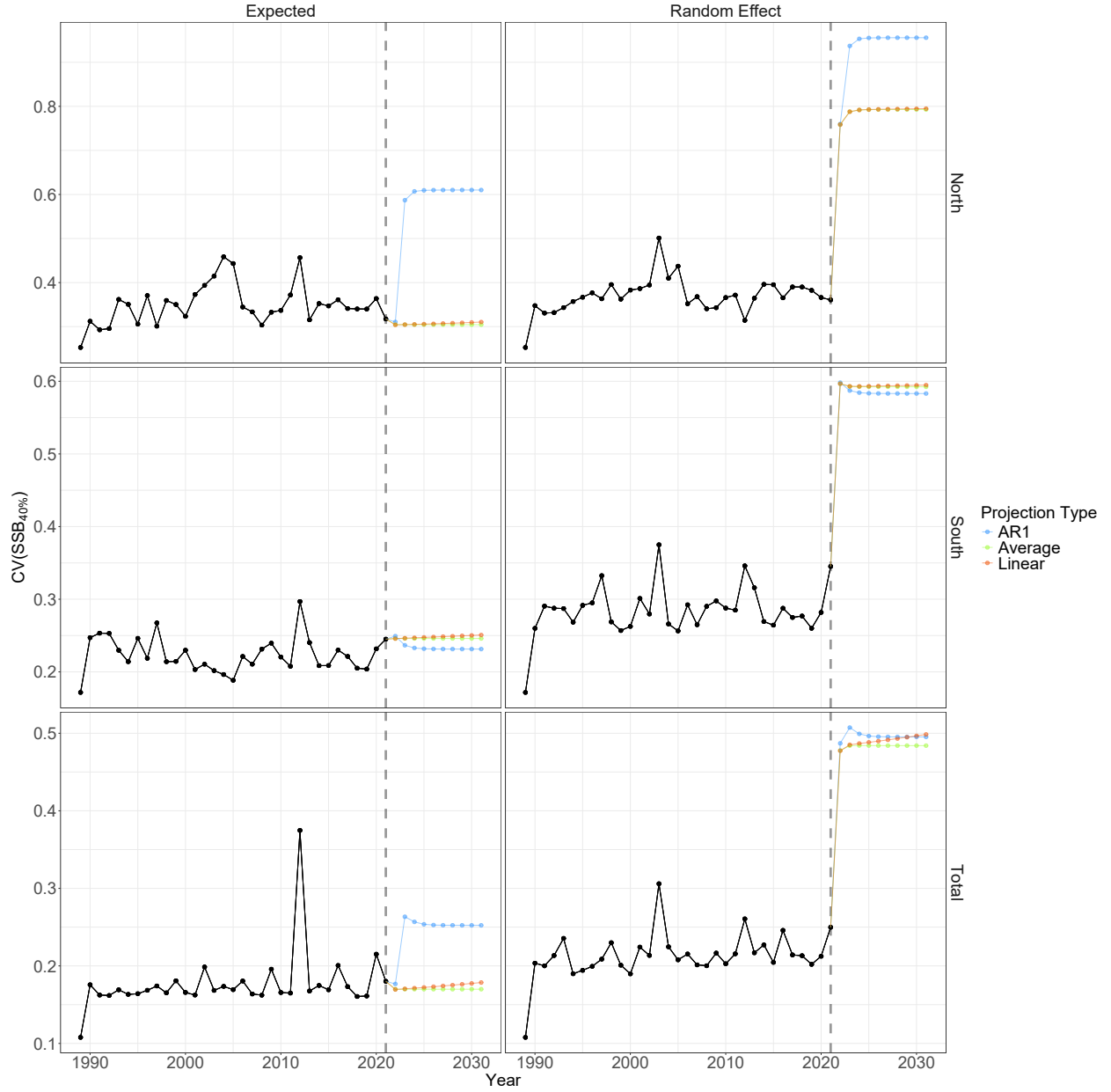


Figure S7: Coefficients of variation for annual equilibrium  $SSB_{40\%}$  as a function of annual expected recruitment or recruitment random effects and annual inputs to  $\phi(\bar{F})$  calculations and alternative annual recruitment types. Estimates in years after 2021 are from projecting model  $M_1$  under three alternative assumptions for the bottom temperature anomalies. Vertical dotted lines indicate the last year of data.

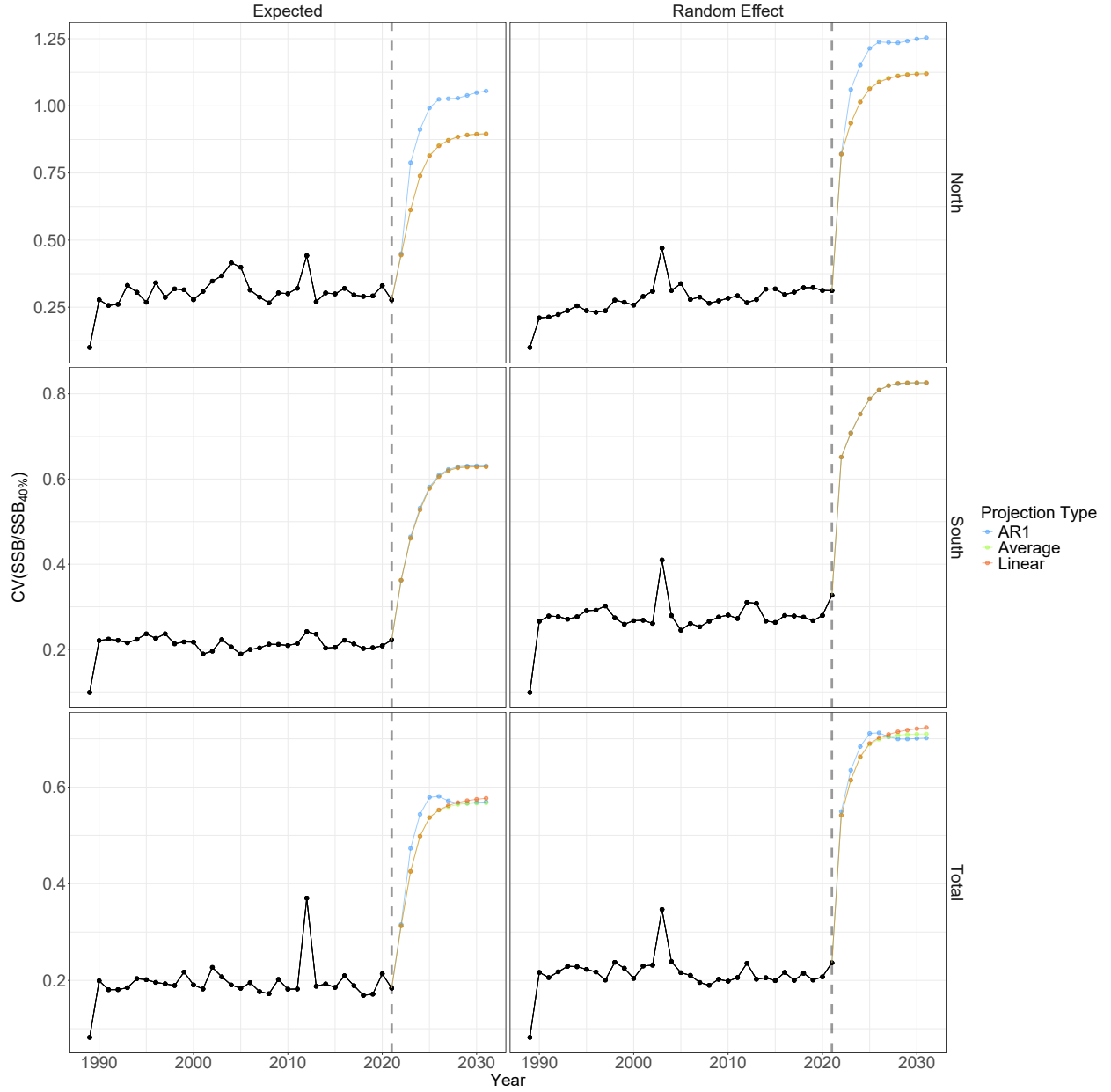


Figure S8: Coefficients of variation for annual ratios of SSB and equilibrium SSB<sub>40%</sub> where the latter is a function of annual expected recruitment or recruitment random effects and annual inputs to  $\phi(\bar{F})$  calculations. Estimates in years after 2021 are from projecting model  $M_1$  under three alternative assumptions for the bottom temperature anomalies. Vertical dotted lines indicate the last year of data.

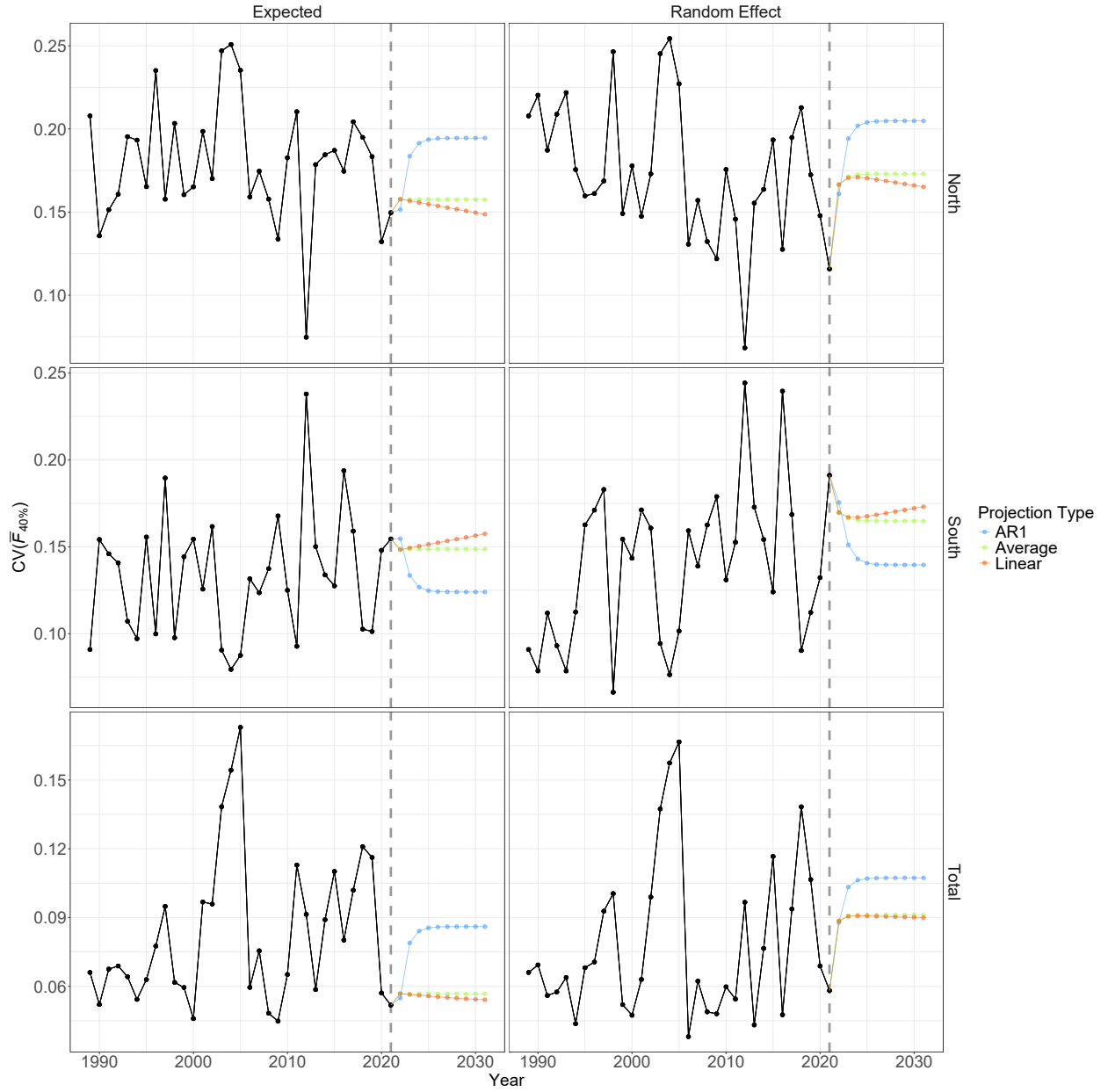


Figure S9: Coefficients of variation for annual equilibrium average  $F$  at ages 6 and 7 that produces the 40% spawning potential ratio as a function of annual expected recruitment or recruitment random effects and annual inputs to  $\phi(\bar{F})$  calculations. Estimates in years after 2021 are from projecting model  $M_1$  under three alternative assumptions for the bottom temperature anomalies. Vertical dotted lines indicate the last year of data.

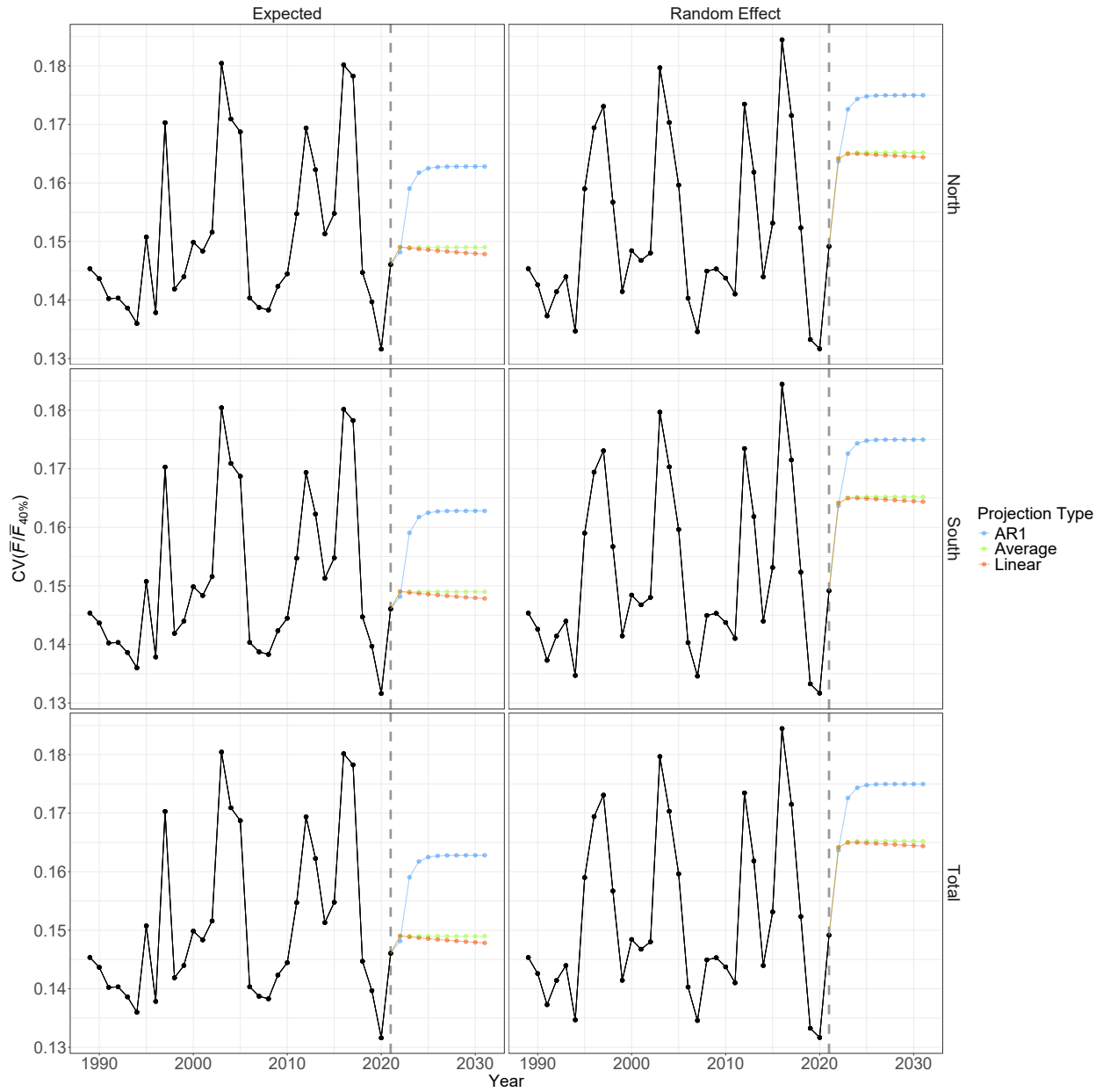


Figure S10: Coefficients of variation for annual ratios of average fishing mortality and equilibrium  $\bar{F}_{40\%}$  at ages 6 and 7 where the latter is a function of annual expected recruitment or recruitment random effects and annual inputs to  $\phi(\bar{F})$  calculations. Estimates in years after 2021 are from projecting model  $M_1$  under three alternative assumptions for the bottom temperature anomalies. Vertical dotted lines indicate the last year of data.

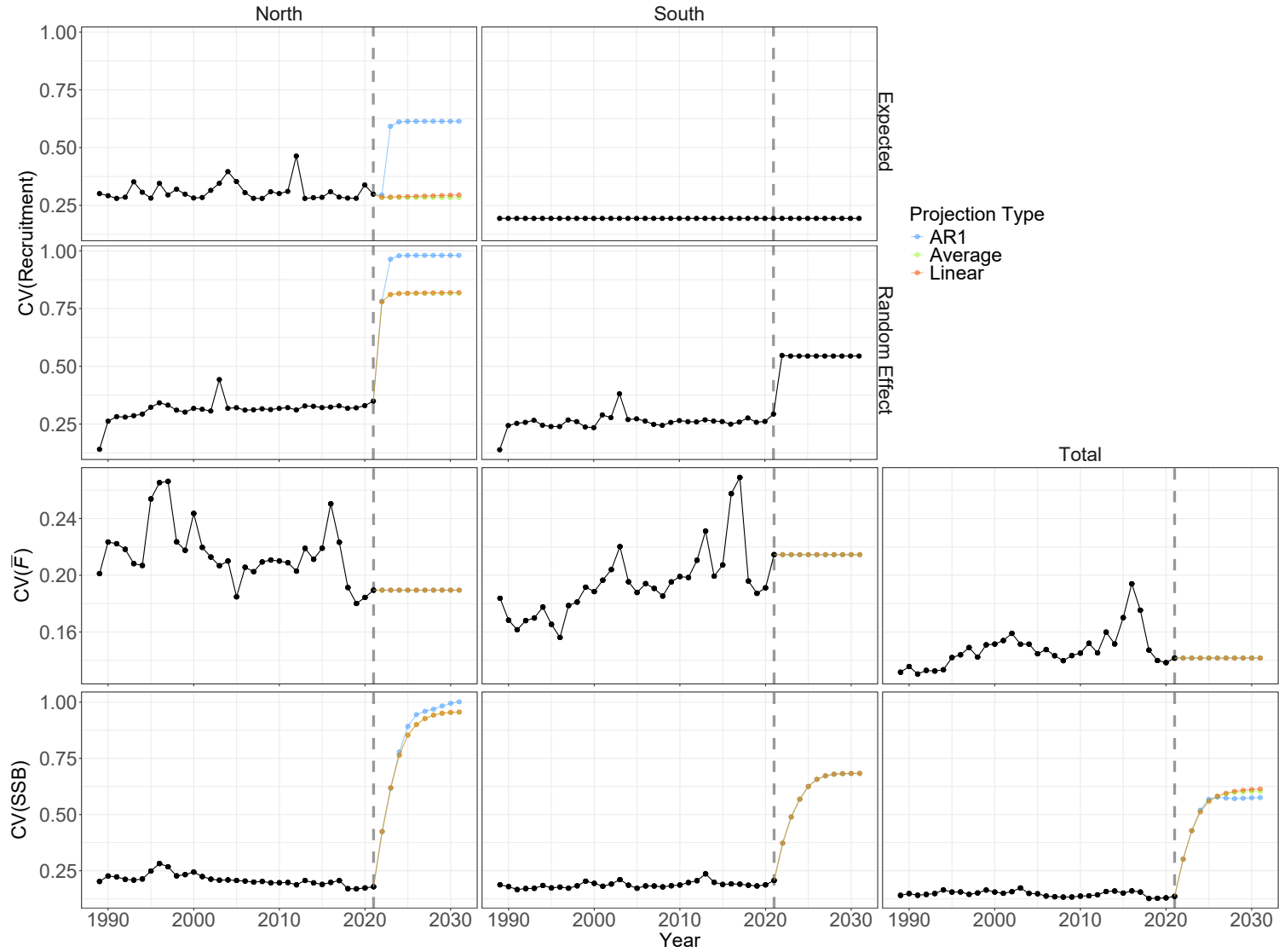


Figure S11: Coefficients of variation for estimates of alternative recruitment estimates (random effects or expected), average fishing mortality at age 6 and 7, and SSB by region and in total from model  $M_1$ . Values in years after 2021 are from projecting model  $M_1$  under three alternative assumptions for the bottom temperature anomalies. Vertical dotted lines indicate the last year of data.



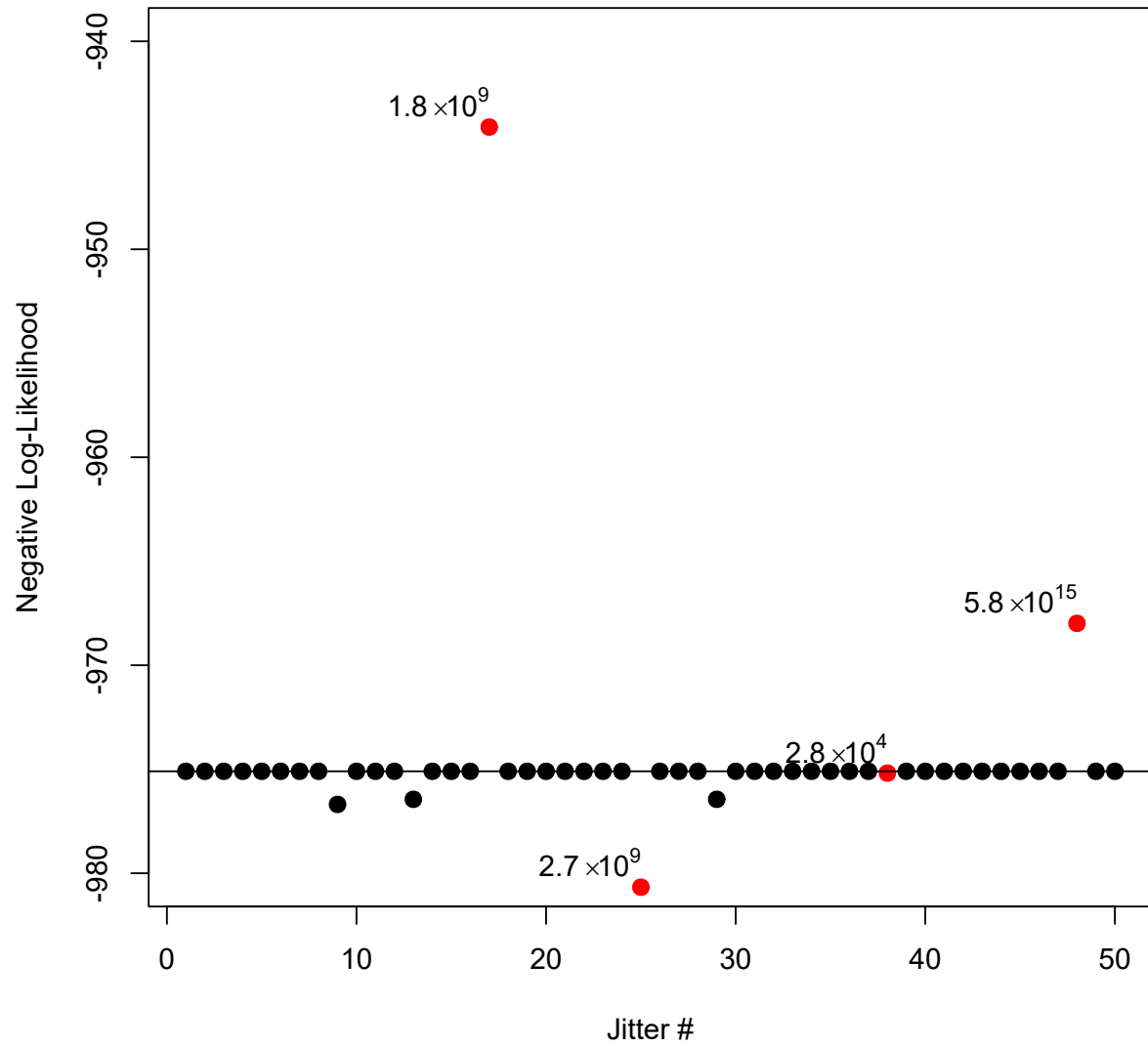


Figure S12: Minimized negative log-likelihood for 50 fits where minimization used initial parameter values jittered from those provided by an initial fit for model  $M_0$ . Black jitters had maximum absolute gradient values  $< 10^{-2}$  and red jitters had values  $> 1$ .

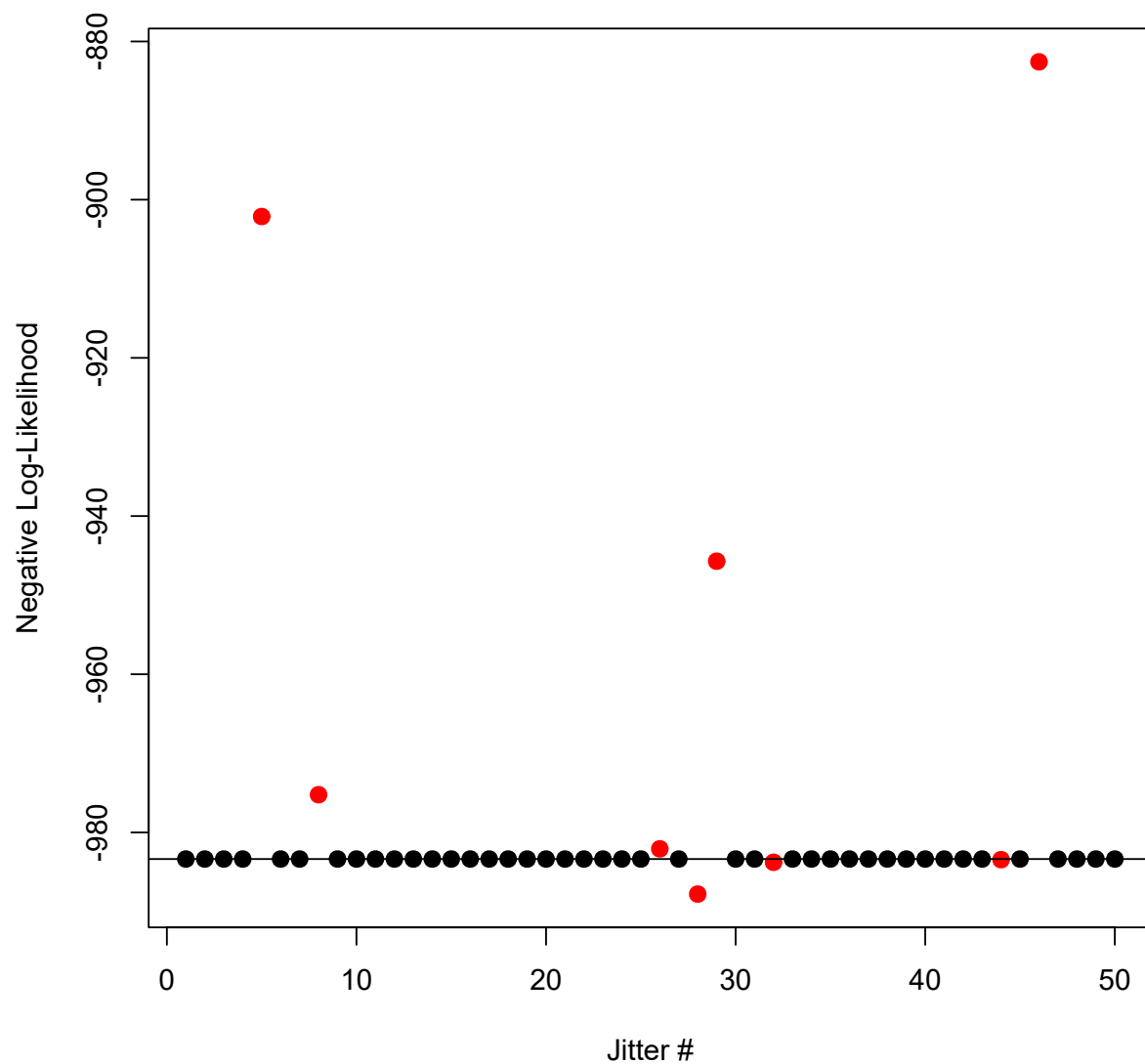


Figure S13: Minimized negative log-likelihood for 50 fits where minimization used initial parameter values jittered from those provided by an initial fit for model  $M_1$ . Fits with black dots had maximum absolute gradient value  $< 0.01$  and fits with red dots had values  $> 10$ .

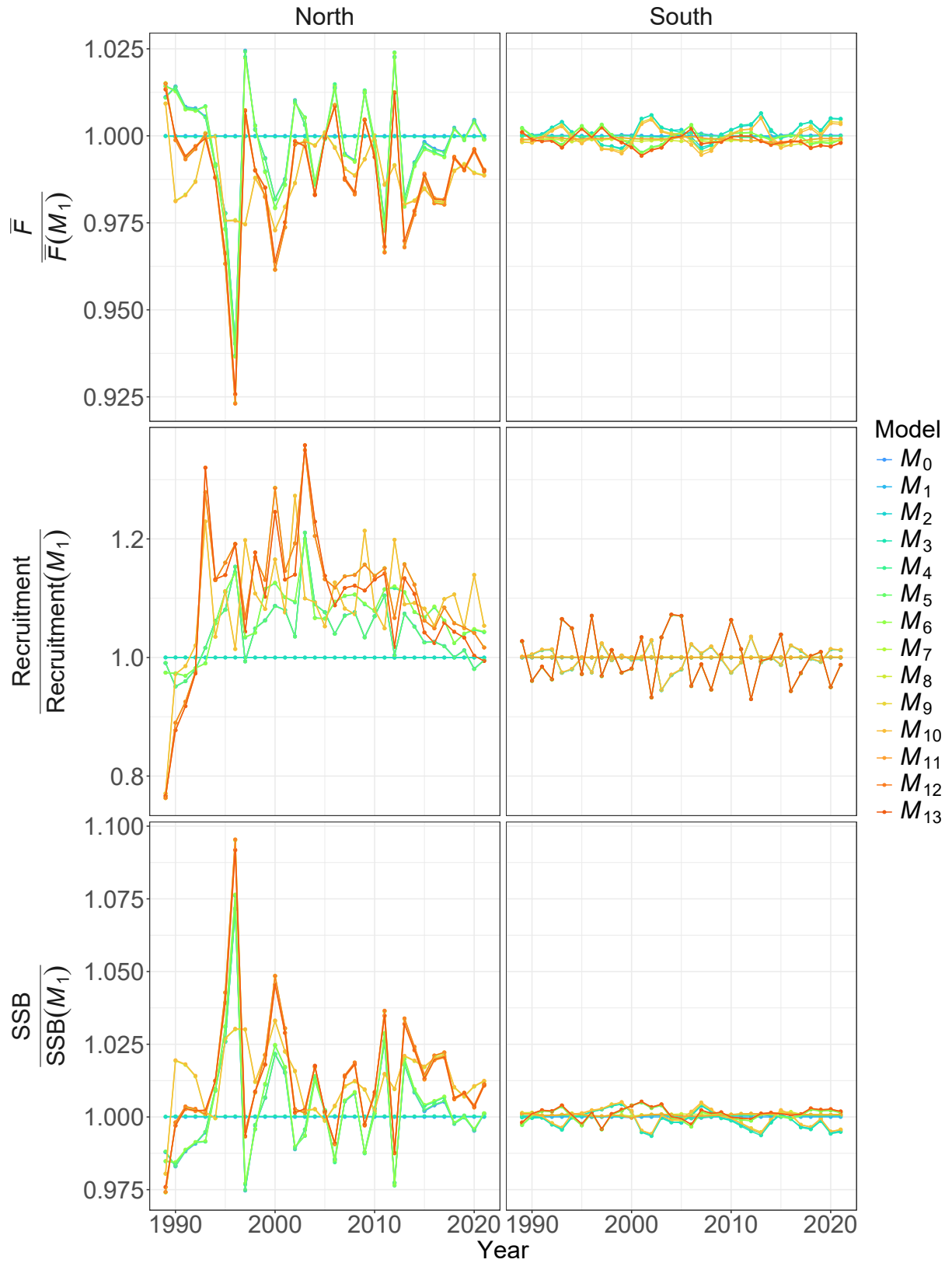


Figure S14: Estimates of SSB,  $F$ , and recruitment relative to those of the best performing model,  $M_1$ .

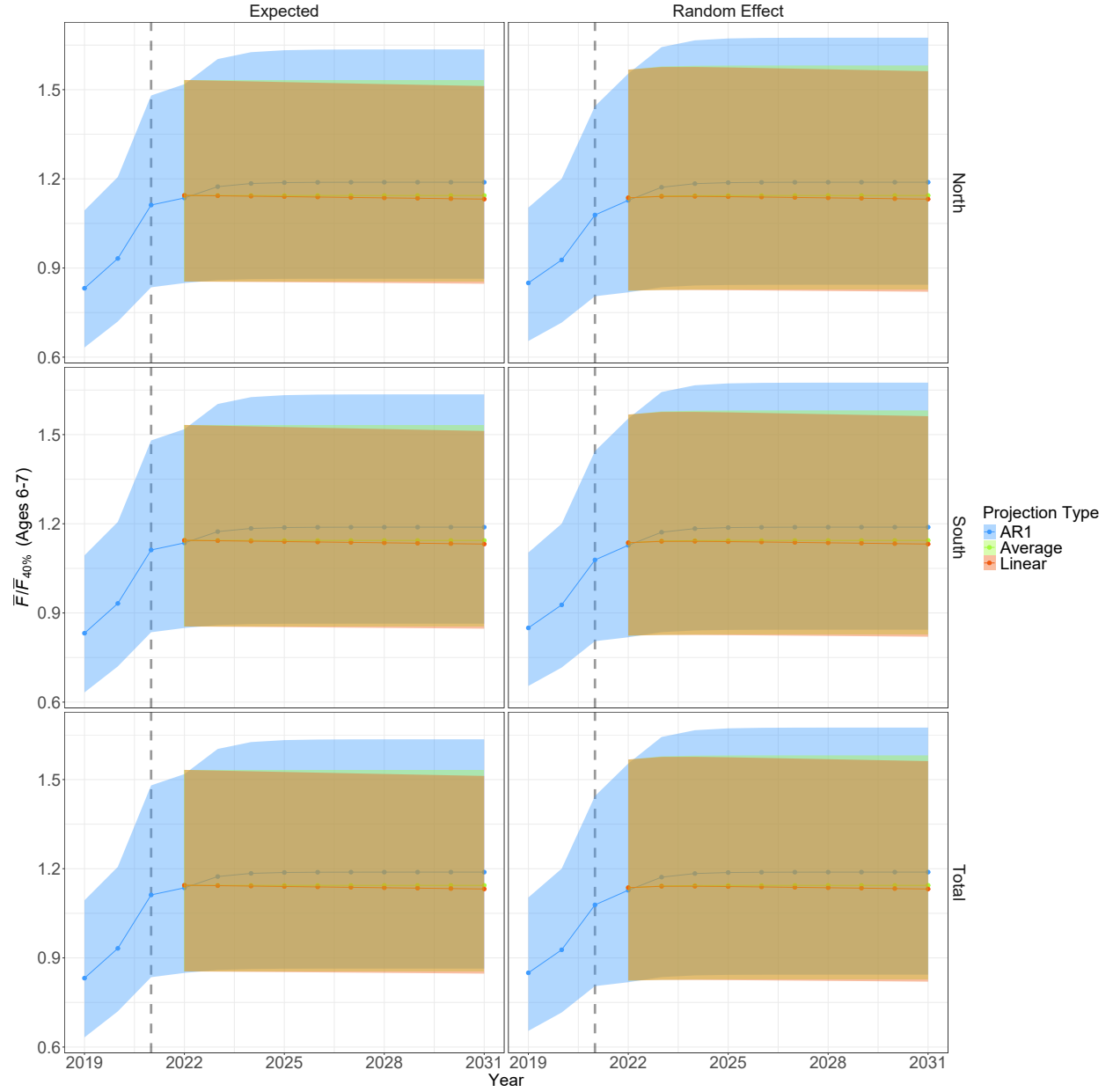


Figure S15: Annual estimates of ratios of fishing mortality to  $F_{40\%}$  by region and in total. Estimates in years beyond 2021 are from projecting model  $M_1$  under alternative assumptions for bottom temperature anomalies in the northern region. Vertical dotted lines indicate the last year of data and polygons represent 95% confidence intervals.

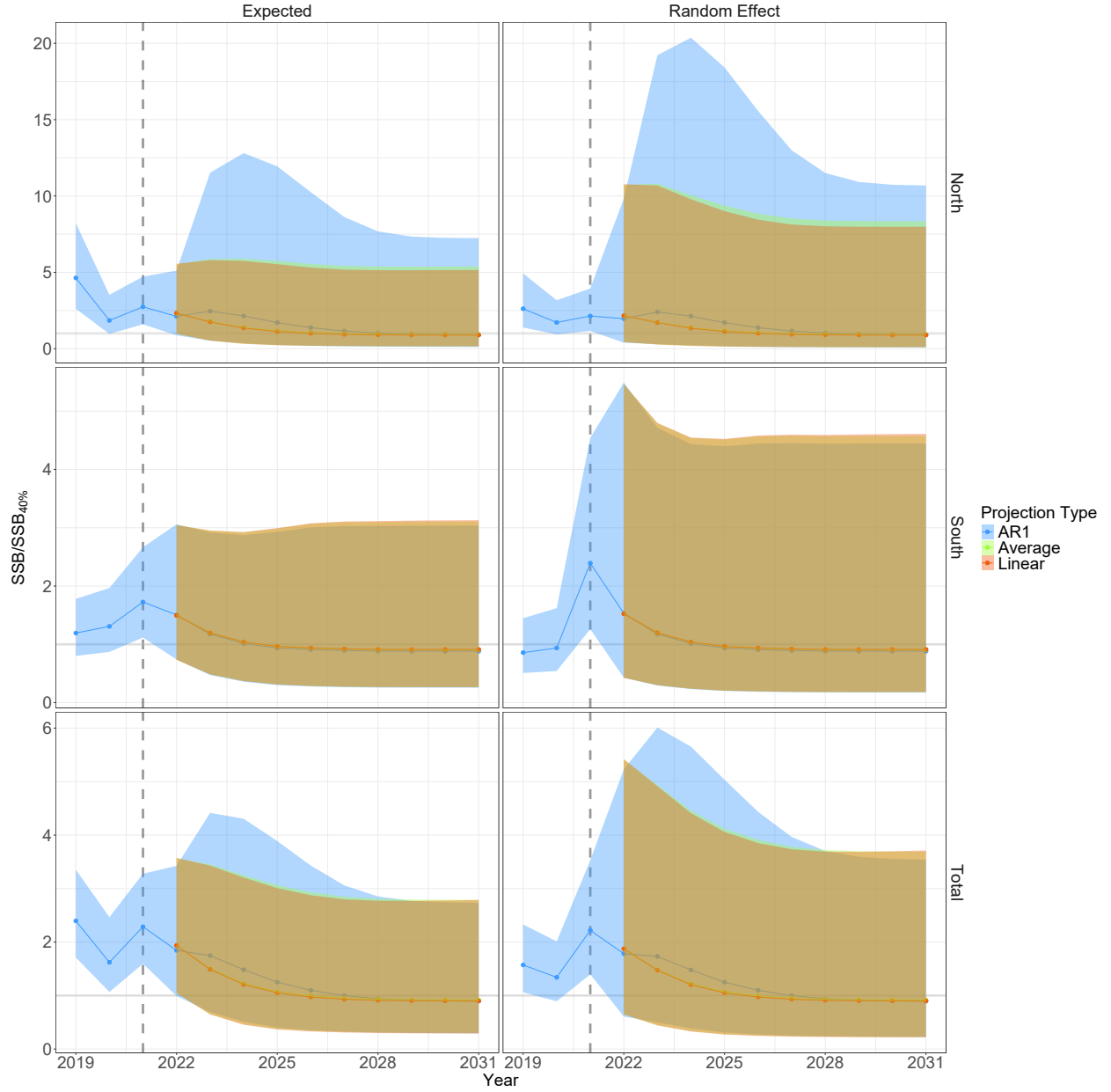


Figure S16: Annual estimates of ratios of SSB to  $SSB_{40\%}$  by region and in total. Estimates in years beyond 2021 are from projecting model  $M_1$  under alternative assumptions for bottom temperature anomalies in the northern region. Vertical dotted lines indicate the last year of data and polygons represent 95% confidence intervals.